DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service

Food and Drug Administration
1401 Rockville Pike
Rockville, MD 20852-1448

CENTER FOR BIOLOGICS EVALUATION AND RESEARCH
Biostatistics Branch (HFM-215)

Memorandum

DATE: *September 5, 1997*

FROM: *Jawahar Tiwari, Ph.D.* JTiwari

THROUGH: *Peter A. Lachenbruch, Ph.D., Chief* P Lachenbruch

SUBJECT: *PLA 96-1408*
REGRANEX® (becaplermin) Gel (recombinant human platelet-derived growth factor) in the treatment of diabetic ulcers, The R.W. Johnson Pharmaceutical Research Institute.

TO: *Kurt Stromberg, M.D.*
Division of Cytokine Biology (HFM-570)

CC: HFM- 99/Document Control Center: PLA 96-1408
HFM-570/Dr. Marzella
HFM-210/Dr. Ellenberg
HFM-210/Chron - File: OP-5.7

This PLA submission consists of a CAPLA and more than 260 volumes of printed material with study reports, line listings of data, and a multitude of posthoc statistical analyses.

In this submission, the claims of efficacy of *REGRANEX®* for the treatment of lower extremity diabetic ulcers are based on cumulative results from four randomized controlled studies (enrolling a total of 922 patients) in which two dose groups (30 µg and 100 µg *REGRANEX®*) and two types of controls (Standard treatment and Vehicle) were compared. The entry and exclusion criteria were very similar in all these studies. The primary endpoint was defined as 100% wound closure (complete healing of the target ulcer) at the end of 20 weeks of therapy. Time to healing and relative ulcer area at endpoint were defined as important secondary endpoints.

Study 90-22120F

This randomized, double-blind, parallel group, vehicle-controlled, multicenter study was designed to assess the efficacy and safety of topically applied becaplermin 30 μg activity/g gel in the treatment of chronic Stage III or IV nonhealing lower extremity cutaneous wounds in subjects with diabetic mellitus. Subjects were randomly assigned to be treated once daily with either becaplermin (61 subjects) or a visually identical vehicle (57 subjects) for a maximum period of 20 weeks, or until the ulcer had completely healed without drainage or the need for a dressing, or until they exited the study as treatment failures. After the screening visit (Visit 1), subjects enrolled into the study started receiving study medication at visit 2 (baseline visit). Subjects visits were weekly for visit 2 through 6, and biweekly after Visit 6. Ulcer measurements were made at each visit and the dosage of study medication adjusted, if necessary.

One hundred eighteen subjects were enrolled at 10 centers, including 61 who received becaplermin 30 μg/g gel and 57 who received vehicle gel (intent-to-treat).

Study 92-22120K

This randomized, double-blind, parallel group, vehicle-controlled, multicenter study was designed to assess the efficacy and safety of topically applied becaplermin gel (30 or 100 μg/g) in the treatment of chronic Stage III or IV nonhealing lower extremity cutaneous wounds in subjects with diabetic mellitus. Subjects were randomly assigned to be treated once daily with either becaplermin 100 μg/g (124 subjects), becaplermin 30 μg/g (132 subjects), or visually identical vehicle (127 subjects). Study medication was to be administered in conjunction with good wound care for a maximum period of 20 weeks, or until the Target Ulcer had completely healed (i.e., re-epithelialized) without drainage or the need for dressing, or until they exited the study as treatment failures. After the screening visit (Visit 1), subjects enrolled into the study started receiving study medication at Visit 2 (baseline visit). Subject visits were weekly for Visits 2 through 6, and every other week after Visit 6. Ulcer measurements were made at each visit and the dosage of study medication adjusted, if necessary.

Efficacy evaluations were primarily based on the comparison of the incidence of 100% wound closure of the Target Ulcer between each becaplermin group and the vehicle group.

Study PDGF-DBFT-001

This was a randomized, parallel group, evaluator (third-party) blinded, multicenter clinical trial with three-month follow-up and open-label extension comparing vehicle gel (NaCMC gel) and standard therapy (wet-to-dry saline dressings applied every 12 hours) in treatment of chronic, Stage III/IV diabetic ulcers of the lower extremity. A becaplermin (100 μg/g) treatment arm was added to enhance subject recruitment and to provide subjects with an opportunity for active get treatment but was not intended for comparison of efficacy variables. Only one skin ulcer, designated as the Target Ulcer, was treated with study gel (vehicle or becaplermin); other (satellite) ulcers could have received standard therapy or other forms of therapy decided upon by the individual investigator. Treatment consisted of twice daily dressing changes, self-administered at 12-hour intervals for up to 20 weeks or until the wound was completely healed, whichever occurred first. Subjects randomized to the vehicle or becaplermin treatment groups applied a predetermined amount of study get to their Target Ulcer once daily during the morning dressing change; the evening

dressing change consisted of wet-to-dry (in practice, wet-to-moist) saline-soaked gauze pads. Subjects who received standard therapy were treated with only the wet-to-dry saline-soaked gauze pads twice daily to their Target Ulcer. All subjects, regardless of treatment group, were required to remain non-weightbearing on the affected limb and underwent surgical debridement of their Target Ulcer prior to study randomization and on as-needed basis throughput the study. Safety and efficacy information were collected during each study visit. Visits were scheduled at weekly intervals up to Visit 6 and then every other week until the end of the study.

The primary efficacy criterion was the incidence of 100% wound closure (the percentage of subjects with completely healed ulcers at endpoint with no dressing or drainage, i.e., Functional Assessment Score of one); other variables were time to 100% wound closure and relative ulcer area at endpoint.

Study PDGF-DBFT-002

The primary objective of this 20-week study was to evaluate the efficacy of becaplermin gel as compared to standard therapy after 20 weeks in the healing of chronic, lower extremity, full-thickness, diabetic ulcers.

This randomized, multicenter, third-party (evaluator) blind, pharmacoeconomic study was designed to compare once-daily topical treatment with becaplermin gel and standard therapy. Study medication was administered in conjunction with good wound care which included twice-daily dressing changes and wound debridement, as needed, to remove necrotic and infected tissue. Subjects had at least one chronic (duration being no less than eight weeks from onset to dosing), lower extremity, full-thickness (Stage III or IV) diabetic ulcer. Subjects were to have a Target Ulcer limb $T_cpO_2 \geq 30$ mmHg at Visit 1 (screening). At Study Week 8, all subjects were assessed as to whether they had responded to treatment. Those who received becaplermin get and who had less than a 30% decrease in their Target Ulcer area were considered poor responders, removed from treatment with becaplermin gel, and placed on standard therapy for remainder of the 20-week treatment phase. The 20-week treatment phase was followed by a 16-week standard therapy phase, during which all subjects who did not heal during the treatment phase received standard therapy only. There was a three-month follow-up for those subjects who healed at any time during the 36-week study to assess ulcer recurrence.

The primary efficacy variable was incidence of 100% closure of the Target Ulcer after 20 weeks of treatment.

COMMENTS

1. This reviewer has compared the results obtained from the data provided in the CAPLA with those given in the study reports and has found no discrepancies. The primary analyses of these studies are consistent with the protocols.

2. The central issue in the evaluation of the efficacy of becaplermin is the consistency and reproducibility of the outcome on the primary endpoint. If the product is efficacious then we should see consistent results in the four well-designed and very similar trials.

2. The **Primary Efficacy Variable: Incidence of 100% Wound Closure in Individual Studies**

The incidence of 100% wound closure in four randomized and controlled studies is given in Table 1.

(a). Efficacy of 30 μg becaplermin: Reproducibility and consistency

(i). In study 90-22120F, 47.5% of the 30 μg becaplermin patients had 100% closure as compared with 24.6% in the control (vehicle) patients. This absolute difference of 22.9% was significant (P=0.013, Table 1).

(ii). Study 92-22120K , a larger study, was also designed to compare the efficacies of 30 and 100 μg becaplermin with the vehicle. Here, the only 36.4% of the 30 μg becaplermin patients had 100% closure as compared with 34.7% in the vehicle-treated patients. The observed absolute difference of 1.8% was not significant.

(iii). Thus, the significant difference observed in study 90-22120F was not confirmed in this trial. However, the becaplermin-treated group showed a higher percentage of wound closure, indicating some degree of consistency between the two studies.

(b). Efficacy of 100 μg becaplermin: Reproducibility and consistency

(i). In study 92-22120K, 49.6% of the 100 μg becaplermin-treated patients showed 100% closure. The efficacy in this high-dose group was significantly different from the vehicle group (an absolute difference of 15.0%, P=0.021). The 30 μg becaplermin group was also significantly different from the 100 μg group (13.2% absolute difference, P=0.043).

(ii). The relatively large PDGF-DBFT-002 trial was primarily designed to compare the efficacy of 100 μg becaplermin with that of the standard treatment. The observed frequencies of 100% wound closure in the standard and 100 μg becaplermin arms were 32.0% and 35.9%, respectively. This difference of 3.9% was not significant (P=0.593). Here again, the statistical significance could not be reproduced. Nevertheless, the percentage of 100% wound closure is higher in the becaplermin arm.

(iii). The PDGF-DBFT-001 trial was designed to compare the standard therapy with the vehicle gel. The 100 μg becaplermin arm was added to study NOT to for comparison of efficacy variables but to enhance subject recruitment. The standard and vehicle arms were found to be not significantly different (22.1% vs 35.7%, P=0.093).

However, in this trial, 100 μg becaplermin arm was significantly better than the standard arm (22.1% vs 44.1%, P=0.037).

(c). Dose-Response in Study 92-22120K

In study 92-2219K, there is a significant evidence of dose response (P=0.017, Jonckheere-Terpstra test) providing some support for the efficacy of becaplermin.

(d). Standard vs Vehicle

As noted in the preceding sections, the PDGF-DBFT-001 study was designed to compare the efficacy of standard treatment with the Vehicle gel. The observed percentages of 100% closure in standard and vehicle arms were 22.1 and 35.7, respectively. This observed difference of 13.6% was not significant (P= 0.093).

(e). Heterogeneity of Response (100% wound closure) Within a Treatment Arm

Since a treatment arm (standard, vehicle, 30 μg, and 100 μg becaplermin) was compared in two or three of the four randomized controlled trials, a comparison of 100% wound closure between the studies reveals an interesting feature of the results. The observed difference between the highest and the lowest percentages in a particular arm are given in Table 2.

These four trials are very similar with respect to inclusion/exclusion criteria, measurement of the primary endpoints, blinding, study sites, dosing and regimen of the treatment, and the conduct of the trial. And, of course, are done by the same sponsor. Thus, any observed difference in a particular arm is largely due to the variability associated with the natural history of the wound and individual patient's attention to the care of his/her conditions. It provides some indication of the background variability ("noise") in the outcome regardless of the type of treatment given to these patients.

The data in Table 2 indicate that this background variability is remarkably similar in all four arms - about 10 to 14%. In the spirit of exploratory analysis, it is interesting to compare the results in a single arm from two different studies with extreme outcomes and see that they are very close to being statistically significant at 0.05 level.

In contrast, the maximum difference between 30 μg becaplermin and vehicle arms was 22.9% (study 90-22120F) and between 100 μg becaplermin and standard was 22.0% (study PDGF-DBFT-001). Thus, in comparison with a "real treatment effect" the amount of background variability is quite large. In this situation, a relatively large trial would be needed to reproduce a statistically significant difference between becaplermin and the control - standard or vehicle - groups (more about this in the COMMENT section of this review).

3. Combined Analysis of Crude Rates

I. The Primary Efficacy Variable: Incidence of 100% Wound Closure in a Combined Analysis of all Four Studies

A comparison of the incidence of 100% wound closure can also be made on the data combined from all four studies. The combined data (total number of patients enrolled and number (percent) with 100% wound closure) is given in the shaded row of Table 1. The results of a statistical analysis of the comparisons between the arms are given in part 2 of Table 1. The difference between the control and the becaplermin arms is clearer in the combined analyses:

(a). **Efficacy of 30 μg becaplermin**:

(i). The 30 μg becaplermin group was significantly different from the standard treatment group (39.9% vs 28.4%, P=0.024).

(ii). The 30 μg becaplermin group was not significantly different from vehicle group (39.9% vs 32.7%, P=0.135).

(b). **Efficacy of 100 μg becaplermin**:

(i). The 100 μg becaplermin group was significantly different from the standard treatment group (42.8% vs 28.4%, P=0.002).

(ii). The 100 μg becaplermin group was also significantly different from the vehicle group (42.8% vs 32.7%, P=0.017).

(c). Dose-Response

There was a significant evidence of dose-response relationship in comparing Standard - 30 μg becaplermin - 100 μg becaplermin groups (P=0.003). A similar significant dose-response relationship was also evident in the comparison of vehicle - 30 μg becaplermin - 100 μg becaplermin (P=0.017). These results provide additional support for the efficacy of becaplermin.

(d). **30 μg vs 100 μg becaplermin**

In the combined data there is no significant difference between 30 μg and 100 μg becaplermin arms (39.9% vs 42.8%, P=0.571).

4. Combined Analysis: Common Odds Ratio

Another approach to combining data from studies with similar control and becaplermin dose arms is to estimate the common Odds Ratio from two or more of the 2 x 2 tables. Each study provides one 2 x 2 table and thus constitutes a stratum.

As the summary data in Table 3 indicate, each of the three comparisons - Vehicle vs 30 μg Regranex, Vehicle vs 100 μg Regranex, and Standard with 100 μg Regranex - was evaluated in two different trials. The estimates of the common Odds Ratios and their 95% Confidence Intervals for these three comparisons are given in the fourth column of the Table.

The estimated common Odds Ratio for the Vehicle vs 100 μg Regranex comparison is 1.7 with the 95% CI from 1.1 to 2.7. The P-value is 0.01. This is the only comparison where 95% CI does not include 1 and the P-value is less than 0.05. The results for the other two comparisons are close to the borderline.

The results of the Homogeneity test are given in the third column of Table 3. None of the P-values is significant at 0.05 level indicating that studies are homogeneous and there is a common Odds Ratio across the strata (studies).

5. Combined Analysis: Average P-value and Summary P-value

There are a number of other approaches to combining data from different trials. Here, we consider two additional simpler methods - averaging P-values and Summary P-values - from a review of methods for combining randomized clinical trials by DeMets (1987) [DeMets DL, Statistics in Medicine, 6:341-348, 1987].

The results from these two methods are given in Table 4. Here again, no comparison gives a P-value of less than 0.05. However, the Summary P-values for the Vehicle vs 100 μg Regranex and Standard vs 100 μg Regranex comparisons are two of the smallest P-vales in this analysis.

CONCLUSION

1. The statistically significant difference between the becaplermin and control groups seen in one study was not reproduced in the subsequent confirmatory trials.

2. However, there is some degree of consistency in the results from these four trials. In all studies percentage of 100% wound closure in the becaplermin groups is higher than that in the control group.

3. In the combined analysis of the crude rates, there is a clear significant difference (P=0.002) between 100 μg becaplermin and the standard treatment groups (42.8% vs 28.4%; a difference of 14.4%). In this analysis, 100 μg becaplermin group is also significantly (P=0.017) different from the vehicle control group 42.8% vs 32.7%; a difference of 10.1%).

4. The design of study 92-22120K incorporated three parallel arms: vehicle, 30 μg, and 100 μg becaplermin. The results from this study show a significant evidence of dose-response relationship. The combined analysis also shows significant evidence of dose-response relationship in Vehicle - 30 μg - 100 μg and Standard - 30 μg - 100 μg comparisons. The existence of a dose response relationship in 92-22120K study and in combined data provides additional support for the efficacy of 100 μg becaplermin.

5. We have used several methods of combining data from these studies. As one would expect, the results are somewhat method dependent. But, in general, they do show a trend in favor of the efficacy of 100 μg Regranex.

6. The lack of reproducibility of statistical significance may be due to relatively small number of patients enrolled in each of the four trials. The results on the primary endpoint in a treatment arm (standard or vehicle or becaplermin) show large variability between the studies (in comparison with the observed difference between becaplermin and the control within a study). Under these conditions, the trial size must be quite large to detect a statistically significant difference between becaplermin and the control arms.

Let us assume that we are interested in designing a new confirmatory trial with 80% power and 5% alpha level. The combined data show that the incidence of 100% wound closure is 28.4% in the standard control arm as compared with 42.8% 100 μg becaplermin arm. We can take a somewhat conservative figure of 30% for the standard and 40% for the becaplermin (generally, in randomized trials patients in the control arm show better results and those in the treatment arms show results less than expected on the basis of limited Phase 1 or 2 data). With these assumptions, a hypothetical trial would need to enroll a total of approximately 800 patients.

In the combined data, we only have a total of 475 patients (190 in the standard and 285 in the 100 μg becaplermin group)! A trial of 500 patients (250/arm) would have only 65% power and, most likely, we would not be able to see a statistically significant difference.

7. Considering the totality of the evidence in the PLA, this reviewer's opinion is that the proverbial glass is 60% full.

Table 1. Incidence of 100% wound closure in four controlled studies.

STUDY	STANDARD		VEHICLE		30 µg Regranex		100 µg Regranex	
	N	*No. with closure(%)*	N	*No. with closure (%)*	N	*No. with closure(%)*	N	*No. with closure (%)*
90-22120F			**57**	14 (24.6)	**61**	29 (47.5)		
92-22120K			**127**	44 (34.7)	**132**	48 (36.4)	**123**	61 (49.6)
PDGF-DBFT-001	**68**	15 (22.1)	**70**	25 (35.7)			**34**	15 (44.1)
PDGF-DBFT-002	**122**	39 (32.0)					**128**	46 (35.9)
ALL 4 Studies Combined	**190**	**54 (28.4)**	**254**	**83 (32.7)**	**193**	**77 (39.9)**	**285**	**122 (42.8)**

COMPARISONS (Individual studies)			**Difference (%)**	**95% CI (%)**	**P***
Vehicle	**vs**	**Standard**			
Study 001			13.6	-1.3, 28.6	0.093
30 µg Regranex	**vs**	**Standard**			
no study					
100 µg Regranex	**vs**	**Standard**			
Study 001			22.0	2.7, 41.4	0.037
Study 002			3.9	-7.8, 15.7	0.593
30 µg Regranex	**vs**	**Vehicle**			
Study F			22.9	6.2, 39.8	0.013
Study K			1.8	-9.9, 13.4	0.796
100 µg Regranex	**vs**	**Vehicle**			
Study K			15.0	2.8, 27.1	0.021
Study 001			8.4	-11.7, 28.5	0.520
100 µg Regranex	**vs**	**30 µg Regranex**			
Study K			13.2	1.2, 25.3	0.043
Dose-response: *vehicle - 30 µg - 100 µg*					
Study K					0.017**

Table 1

Part 2 (Combined Analysis)

STUDY	STANDARD		VEHICLE		30 µg Regranex		100 µg Regranex	
	N	*No. with closure(%)*	N	*No. with closure (%)*	N	*No. with closure(%)*	N	*No. with closure (%)*
ALL 4 Studies Combined	**190**	**54 (28.4)**	**254**	**83 (32.7)**	**193**	**77 (39.9)**	**285**	**122 (42.8)**

COMPARISONS (All 4 studies combined)			Difference (%)	95% CI (%)	P*
Vehicle	vs	Standard	4.3	-4.4, 12.9	0.352
30 µg Regranex	vs	Standard	11.5	2.0, 20.9	0.024
100 µg Regranex	vs	Standard	14.4	5.8, 23.0	0.002
30 µg Regranex	vs	Vehicle	7.2	-1.8, 16.2	0.135
100 µg Regranex	vs	Vehicle	10.1	2.0, 18.3	0.017
100 µg Regranex	vs	30 µg Regranex	2.9	-6.1, 11.9	0.571

Dose-response:

Standard - 30 µg - 100 µg	0.003**
Vehicle - 30 µg - 100 µg	0.017**

* Fisher's Exact test
** Jonckheere-Terpstra test

Table 2. The observed difference in 100% wound closure between two studies with extreme results in a treatment arm.

TREATMENT ARM	STUDIES WITH EXTREME RESULTS	DIFFERENCE (95% CI)	P-VALUE
Standard	PDGF-DBFT-002 **(32.0%)** VS PDGF-DBFT-001 **(22.1%)**	9.9% (-3.0, 22.8)	0.18
Vehicle	PDGF-DBFT-001 **(35.7%)** VS 90-22120F **(24.6%)**	11.1% (-4.7, 27.0)	0.25
30µg Regranex	90-22120F **(47.5%)** VS 92-22120K **(36.4%)**	11.1% (-3.8, 26.2)	0.16
100 µg Regranex	92-22120K **(49.6%)** VS PDGF-DBFT-002 **(35.9%)**	13.7% (1.5, 25.8)	0.03

Table 3. Estimates of Common Odds Ratio and the associated 95% Confidence intervals.

Comparison	Studies included in this analysis	Heterogeneity P-value	Common Odds Ratio (95% CI)	P-value
Vehicle vs 30μg Regranex	90-22120F 92-22120K	0.05	1.4 (0.9, 2.2)	0.09
Vehicle vs 100μg Regranex	92-22120K PDGF-DBFT-001	0.6	1.7 (1.1, 2.7)	0.01
Standard vs 100μg Regranex	PDGF-DBFT-001 PDGF-DBFT-002	0.1	1.5 (0.9, 2.3)	0.09

Table 4. Some other methods of combined analysis*

COMPARISON	STUDY	Observed P-value	METHOD	
			Averaging P-value	Summary P-value
30 ug vs Vehicle	F	0.013	0.413	0.327
	K	0.796		
100 ug vs Vehicle	K	0.021	0.416	0.146
	DBFT-001	0.520		
100 ug vs Standard	K	0.037	0.506	0.199
	DBFT-002	0.593		

*** DeMets DL, Statistics in Medicine, 6:341-348, 1987**